

大模型单元测试用例自动生成分析报告

11 组

一、 背景

使用 ChatGPT 3.5 进行测试。

发送项目架构和主要代码，要求生成下列文件单元测试。

聚焦测试结果和覆盖率，进行测试质量分析。

二、 单元测试

提示词：粘贴源代码，要求生成单元测试

1、 Console ChatGPT

结果：共生成 7 个单元测试，测试有 3 个失败，4 个成功。class 覆盖率 50%，method 42%，line 覆盖率 18%。

分析：失败率较高，主要是模型不能很好读取整个项目的工作原理，因此代码运行逻辑出现问题。给出错误信息修改后，能够运行，但是测试本身意义不明。

2、 ChatGPT Stream

结果：共生成 5 个单元测试，测试有 1 个失败，4 个成功，class 覆盖率 100%，method 覆盖率 100%，line 覆盖率 93%。

分析：代码比较简单，测试效果良好。

3、 ChatGPT

结果：共生成 9 个单元测试，测试有 1 个失败，8 个成功，class 覆盖率 100%，method 覆盖率 41%，line 覆盖率 37%。

分析：由于代码较长，模型很难为所有方法同时生成测试，所以覆盖率相当低，有相当多的情况没有写测试。

4、 Chat Context Holder

结果：共生成 5 个单元测试，测试有 0 个失败，5 个成功，class 覆盖率 100%，method 覆盖率 100%，line 覆盖率 100%。

分析：代码比较简单，测试效果优秀。

5、 Format Date Util

结果：共生成 2 个单元测试，测试有 0 个失败，2 个成功，class 覆盖率 100%，method 覆盖率 100%，line 覆盖率 100%。

分析：代码比较简单，测试效果优秀。

6、 Proxys

结果：共生成 2 个单元测试，测试有 0 个失败，2 个成功，class 覆盖率 100%，method 覆盖率 100%，line 覆盖率 100%。

分析：代码比较简单，测试效果优秀。

7、 See Helper

结果：共生成 4 个单元测试，测试有 0 个失败，4 个成功，class

覆盖率 100%，method 覆盖率 100%，line 覆盖率 100%。

分析：代码比较简单，测试效果优秀。

8、 Tokens Util

结果：共生成 10 个单元测试，测试有 0 个失败，10 个成功，

class 覆盖率 100%，method 覆盖率 100%，line 覆盖率 100%。

分析：代码比较简单，测试效果优秀。

9、 Abstract Stream Listener

结果：共生成 4 个单元测试，测试有 1 个失败，3 个成功，class

覆盖率 100%，method 覆盖率 50%，line 覆盖率 77%。

分析：代码虽然比较简单，但模型很难完全记住项目结构，导致给出的源代码就是有问题的，经过修改后可以运行，但是并不全对。

10、 Console Stream Listener

结果：共生成 4 个单元测试，测试有 0 个失败，4 个成功，class

覆盖率 100%，method 覆盖率 100%，line 覆盖率 100%。

分析：代码比较简单，测试效果优秀。

11、 See Stream Listener

结果：共生成 2 个单元测试，测试有 0 个失败，2 个成功，class 覆盖率 100%，method 覆盖率 100%，line 覆盖率 100%。

分析：代码比较简单，测试效果优秀。

总测试结果如下：

